DOCUMENT RESUME

ED 239 428 EC 161 165

AUTHOR Hicks, John S.
TITLE Validation Studies of the Five P's. The Five P's: A New Handicapped Preschool Children's Assessment Tool.
PUB DATE Aug 83
NOTE 24p.; Paper presented at the Annual Convention of the American Psychological Association (91st, Anaheim, CA, August 26-30,1983).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Behavior Rating Scales; *Classroom Observation Techniques; *Disabilities; Evaluation Methods; Preschool Education; Student Evaluation; *Test Validity
IDENTIFIERS *Parent Professional Preschool Performance Profile

ABSTRACT

This paper reports the results of the validity studies done on the Five P's (Parent/Professional Preschool Performance Profile), an assessment tool for parents and teachers developed by the Variety Preschooler's Workshop (Syosset, New York) to rate young handicapped children functioning between birth and 5 years of age on their observed performance in five areas of development: (1) routines and self-help skills, (2) motor, (3) language, (4) social, and (5) cognitive development. The validity studies focused on the correlations of the teacher's rating of 14 of the Five P's scales with scores on established tests including selected subscales of the Preschool Attainment Record; California Preschool Social Comptency Scale; the Developmental Test of Visual Motor Integration; subscales of the McCarthy Scale; subscales of the Burkes Behavior Rating Scale; the Carrow Test for the Auditory Comprehension of Language; the Bzoch-League REEL scale; a Mean Length of Utterance measure; and the Leiter scale. The validity coefficients for all 14 subscales were in the .50 to .70 range. The results indicate strong evidence that the Five P's is a valid teacher observation instrument which measures important aspects of the young handicapped child's behaviors and skills. (Author/CL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED239428

The Five P's

A New Handicapped Preschool Children's Assessment Tool

John S. Hicks, Ed.D., Fordham University, New York, N. Y.

Validation Studies of the Five P's

Paper Presented at 91st Annual Convention of the
American Psychological Association at Anaheim, California, August 1983

EC161165

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
John S. Hicks
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

VARIETY PRE-SCHOOLER'S WORKSHOP
47 Humphrey Drive, Syosset, N.Y. 11791

Abstract

Validation Studies of the Five P's

by John S. Hicks, EdD.

This paper reports the results of the validity studies done on the Five P's (Parent/Professional Preschool Performance Profile), developed by Judith S. Bloch, director/founder of Variety Pre-Schooler's Workshop. The Five P's is an assessment tool for parents and teachers to rate young handicapped children functioning between birth and five years of age on their observed performance in five areas of development: routines and self-help skills, motor, language, social and cognitive development. The validity studies focused on the correlations of the teacher's ratings of fourteen of the Five P's scales with scores on established tests including: selected subscales of the PAR; California Preschool Social Competency Scale; the Developmental Test of Visual Motor Integration; subscales of the McCarthy Scales; subscales of the Burks Behavior Rating Scale; the Carrow Test for the Auditory Comprehension of Language; the Bzoch-League REEL scale; a Mean Length of Utterance; and the Leiter. The validity coefficients for all fourteen subscales were in the .50 to .70 range. The results indicate strong evidence that the Five P's is a valid teacher observation instrument which measures important aspects of the young handicapped child's behaviors and skills.

VARIETY PRE-SCHOOLER'S WORKSHOP
47 Humphrey Drive, Syosset, NY 11791

Validation Studies of the Five P's

By: John S. Hicks, Ed.D.

Paper: Presented at 91st Annual Convention of the American Psychological Association at Anaheim, California, August 1983

For the past year, May of 1982 through June of 1983, several studies on the Five P's have been in progress. These were designed to yield evidence of both the reliability and validity of this new teacher/parent observational scale. The evidence relating to the reliability of the instrument was prepared for presentation at another professional conference, but is available for interested parties upon request. This presentation to members of the American Psychological Association will focus on the evidence which relates to the validation studies which were completed during the past year.

Before reviewing this year's results, let me briefly summarize results of the initial studies and validity which were reported in the (fall of 1982) Interim Technical Manual. At that point in time validation was primarily studied using factor analysis (Harman, 1976). Each of the five scales was factor analyzed, using a sample of 172 cases from Pre-Schooler's Workshop. Each of the scales has from 58 to 78 items. The rule for factor analysis recommends a sample of five times the number of variables (Nunnally, 1978, and others). However, it was

thought to be important to look at the solutions which were produced by the factor analyses. Each of the five scales produced solutions, using oblique rotations which represented logically intact constructs -- constructs that seemed to have "clinical" meaning.

As of June of 1983 the data pool had grown from 172 cases to slightly over five hundred cases. This new pool represented ratings from one other school in the New York area plus ratings from the PSW children. At this point the sample clearly represents enough children to satisfy the five for one rule. Efforts are presently being made to broaden further the geographical base of the standardization sample. A second set of factor analyses during the fall of 1983 should yield further substantiating evidence concerning the "construct" validity of each of the five scales.

It was hypothesized that if the items were really measuring important aspects of the children's pathology, then scores on the sub-scales would not correlate highly with the ages of the children. When these correlations were computed, it was found that only 13.0% of the items correlated above $\pm$ .30 with the children's chronological age as calculated in months. Clearly, the items are not age related, at least with our population of

developmentally disabled children, which is exactly the kind of evidence we had hoped for. Expressed another way, it is clear that the items on the Five P's are not measuring typical age dependent behavior; they seem to be measuring skills which are central to the pathology in the children. The assumption is made that pathology is not highly correlated with age, and consequently the scales show some strong degree of validity.

The evidence concerning the validity of the Five P's which was available after the 1981/82 school year studies was positive and encouraging. Initial factor analyses had suggested that the scale was tapping some clinical constructs, and the correlations with age suggested that the behaviors being measured had content validity in as much as they were not just measuring age related behaviors. However, it was understood that several other important aspects of the validity of the five subscales needed to be investigated. Consequently, several validity studies were planned for the fall of 1982, which have now been completed. These studies focused on establishing the "concurrent" validity of the instrument.

Procedures:

During the late summer of 1982, the research team met to plan the further validation study. Each of the five scales:

Routines and Self-Help; Motor Development; Language Development; Social Development and Cognitive Development were constructed of several sub-parts. These sub-parts and the items of each of the five major scales were grouped together from professional judgement over a period of several years when the scale was being developed. Consequently, some of its major scales had an internal structure which had been already established (such as Language Development which had within it receptive language, expressive language and communicative competence). The factor analysis also suggested several constructs which were operating within the scales.

An effort was made to identify an "external" test, instrument, or scale which would measure constructs which we felt were operating in the Five P's. This follows the classical model of establishing validity by comparing scores from a new instrument with scores on an already established instrument with good validity. Whenever possible, efforts were made to identify instruments which would <u>not</u> be used by teachers. That is, in attempting to establish the validity of the Language Development scale on the Five P's, which is a teacher rating scale, instruments (such as the REEL) which the speech therapists at the agency used, were identified so that the judgment of the teacher would be compared

with that of a speech therapist or the school's psychologist.

However, in several cases, the only appropriate match to the teacher judgements built into the scales of the Five P's, were other published scales that basically were designed to be teacher rating scales. As that happened, it was understood that at times a child's teacher was rating the child using both the Five P's and another already established teacher rating scale. In those cases, which will be identified, the correlations reported are the child's teacher rating the child on two similar scales. It is hoped, in the case where no reasonable scale was found which could be applied by another professional who was not the classroom teacher, that the subsequent factor analysis done in the summer of 1983 would provide enough strong evidence to establish the sub-scale organization of each of the five major scales.

Sample:

In establishing a sample to test the validity of each of the five sub-scales, children were randomly selected from each of the fourteen classes in the program at PSW as of October of 1982. Since the class groupings are usually dependent on age and type of problems, the sample approximated a hypothetically stratified sample of the children. Table 1 provides information on the ages of the forty-one children chosen for the validity studies.

The age ratio and the male-female ratio of the children in the validations sample approximated the distribution in the PSW population at large.

Table 1 - Age + Sex of Children in Validation Sample

Age		Frequencies		
Months	Years	Male	Female	TOTAL
18-23	1½	1	0	1
24-29	2	0	1	1
30-35	2½	2	1	3
36-41	3	3	1	4
42-47	3½	5	1	6
48-53	4	3	2	5
54-59	4½	4	2	6
60-65	5	2	2	4
66-71	5½	0	0	0
72-77	6	6	2	8
78-83	6½	3	0	3
	TOTALS	29	12	41

The 41 children who were selected as the validation sample came from a total population of about 85 children as of October 1982. As of June 1983, the school population approximated 115 children. While the 41 children were chosen, they were not identified immediately to the teachers in their classes. The teachers in the classrooms did the "usual" fall testing of all the children in their classes (usually 5 - 7 children). Immediately after the fall testing had been done by the teachers, the validation materials were distributed to the school psychologists, speech therapists, and the teachers. All of the validation ratings by the classroom teachers, speech people and psychologists were completed within a two month period of time. It should be noted that at times the differences in the ages of the children made it necessary to use two different instruments for one construct. Thus, several of the correlation coefficients reported are for samples which are smaller than 41, since our sample spread in age outside the limits of the standardized tests used, and it seemed meaningless to try to project a score when the age limits were exceeded. The following sections present the validity correlations for each of the sub-sections of the five scales.

Results of the Validity Studies:

One further note must be made at this time. The data presented in this paper focuses on the "above the line" teacher

judgements. It does not include any analysis of parent ratings of the children, or of the Interfering Behaviors. These studies were thought to be basically an effort to validate the teacher judgements of children's behavior in their classrooms.

Scale 1 - Routines and Self-Help:

This scale contains a total of 78 items in 4 sub-scales: 27 in classroom adjustment, 15 in toileting and personal hygiene, 15 in mealtime behaviors, and 21 items in dressing. It basically describes some developmental tasks which would appear to be typical in normal children but which seem to present developmental problems for these children. The "external" scales which seemed to be closest to the skills being measured by this scale were the Preschool Attainment Record (Doll, 1966), and the California Preschool Social Competency Scale (Levine et al. 1969).

After the September 1982 ratings by the teachers of the children in the validation sample, those teachers were given both the Preschool Attainment Record and the California Preschool Social Competency Scale and were asked to apply those scales to the children in the sample. Table 2 presents the validity coefficients for each of the 4 sub-scales of Routines and Self-Help and the external tests.

Table 2 - Validity Correlation Coefficients for Routines and Self-Help Sub-scales on the Five P's

External Scales	Classroom Adjustment (1)	N	Toileting/ Hygiene	N	Mealtime Behaviors	N	Dressing Skills	N
1. California-PSCS	.73	39	.62	39	.70	39	.69	39
2. Preschool Attainment Record								
A. Manipulation	.64	41	.61	41	.67	41	.70	41
B. Responsibility	.72	41	.56	41	.65	41	.61	41

(1) All correlations reported in Table 2 were significant at the .01 level.

The evidence from Table 2 above suggests that the four sub-scales: Classroom Adjustment; Toileting/Hygiene; Mealtime Behaviors; and Dressing Skills present strong validity coefficients. This suggests that the teacher ratings on this part of the Five P's is substantially similar to teacher ratings on both the California PSCS and the Preschool Attainment Record. Due to the strong level of the validity coefficients, no corrections for attenuation were calculated. Gilford (1965) and others suggest that in almost all cases validity coefficients are increased when corrections for attenuation are made.

Scale 2 - Motor Development

The Motor Development Scale consists of two sub-scales. The first sub-scale is Gross Motor Skills which includes 32 items. The second sub-scale is Fine Motor Skills which includes 45 items. Scores for each scale were calculated by adding points earned, and a total score was constructed by adding the two sub-scales together.

The Motor Development Scale was validated against four different "external" tests; the Developmental Test of Visual Motor Integration by Beery (1967), sub-scales on the McCarthy (1972), sub-scales on the Burks Behavior Rating Scale (1981), and sub-scales of the Preschool Attainment Record (1966). The VMI and the McCarthy were administered by the School Psychologist on an individual basis while the Burks and the PAR were used by the classroom teachers as is common practice. Table 3 presents the evidence concerning the validity coefficients for the Motor Development Scale.

Table 3 - Validity Correlation Coefficients for the Motor Development Scale Sub-Scales on the Five P's

External Scales	Gross Motor (1)	N	Fine Motor	N	Total Motor Development	N
1. VMI - Beery	.61	28	.72	28	.71	28
2. McCarthy						
A. Leg Coordination	.59	32	.59	32	.63	32
B. Arm Coordination	.51	33	.50	33	.53	33
C. Imitative Action	.73	29	.78	29	.79	29
3. PAR-Manipulation	.60	41	.70	41	.69	41
4. Burks-Poor Coordination	-.66	41	-.56	41	-.62	41

(1) All correlations reported in Table 3 were significant at the .01 level.

The evidence presented in Table 3 suggests strongly that the classroom teacher ratings on both gross motor and fine motor skills correlate strongly with corresponding judgments made by the school psychologists. It also suggested strong agreement between the teacher judgments on the Five P's, the PAR and the Burks. Again, since the correlations were sufficiently strong, no corrections for attenuation were calculated.

Table 3 presents very strong supportive evidence that the Motor Scales contain within them strong elements of validity, when compared with already established instruments such as the VMI or the McCarthy scales.

Scale 3 - Language Development:

The Language Development scale consists of three sub-scales. The first sub-scale is Communicative Competence which includes 20 items, the second sub-scale is Receptive Language which includes 22 items, and the last sub-scale is Expressive Language which includes 24 items. Scores for each of the sub-scales were calculated by adding points earned on each item, and a total language score was constructed by summing the scores on all three sub-scales.

The Language Development scale was validated against several different measures which were applied by the speech therapists in the agency who knew the children best. These measures included a Mean Length of utterance score, the Bzoch-League Receptive-Expressive Emergent Language Scale (1970) and the Carrow Test for the Auditory Comprehension of Language (1973). Table 4 presents the evidence relating to the validity coefficients for the Language Scale.

Table 4 - Validity Correlation Coefficients for the Language Development Scale Sub-Scales on the Five P's

External Scales	Communicative Competence (1)	N	Receptive Language	N	Expressive Language	N	Total Language	N
1. M.L.U.	.62	34	.63	34	.75	34	.71	34
2. REEL-Receptive	.49	21	.64	21	.65	21	.68	21
3. REEL-Expressive	.58	21	.59	21	.78	21	.76	21
4. Carrow	.53	20	.42*	20	.51*	20	.53	20

(1) *All coefficients were significant at the .01 level except for two which were significant at the .05 level and are followed by an asterisk.

The evidence presented in Table 4 suggests a very strong degree of support for the Expressive Language sub-scale, which correlated .75 with the Mean Length of Utterance and .78 with the Expressive scale of the REEL. The Receptive Language sub-scale correlated .64 with the REEL Receptive scale and .42 with the Carrow which is basically a receptive language scale. The Communicative Competence sub-scale correlates moderately highly with all four measures. Communicative Competence, as defined by most language professionals, is an emerging construct which does not have any clearly established "external" tests against which it can be validated. It is interesting to note that the Communicative

Competence sub-scale correlated very highly with several of the other teacher rating scales, notably the ideation sub-scale in the PAR (.70) and the California PSCS (.73). In general, there is very strong support of the Expressive Language sub-scale, moderate support for the Receptive Language and some emerging patterns related to the construct of Communicative Competence.

Scale 4 - Social Development:

The Social Development Scale consists of three sub-scales. The first sub-scale is Emerging Self which consists of 27 items, the second sub-scale is Relationship to Adults which contains 19 items, and the third sub-scale is Relationship to Children which has 23 items. The total scale has 69 items. Scores for each of the sub-scales were calculated by adding points earned on each item, and a total social development score was constructed by summing the scores of all three sub-scales.

The Social Development Scale was validated against three other teacher rating scales, the California Preschool Social Competency Scale, several sub-scales of the Preschool Attainment Record, and several sub-scales of the Burks Behavior Rating Scale. Table 5 presents the evidence relating to the validity coefficients for the Social Development Scale.

Table 5 - Validity Correlation Coefficients

for the Social Development Scale

(1)

Sub-scales on the Five P's

External Scales	Emerging Self	N	Relations to Adults	N	Relations to Children	N	Total Score	N
1. California PSCS	.62	39	.60	39	.79	39	.77	39
2. PAR - Rapport	.32 (.02)	41	.59	41	.69	41	.60	41
3. PAR - Responsibilities	.39	41	.53	41	.62	41	.58	41
4. Burks - Poor Impulse Control	-.41	39	-.34	39	-.20 (NS)	39	-.35	39
5. Burks - Poor Anger Control	-.24 (NS)	39	-.18 (NS)	39	-.08 (NS)	39	-.19(NS)	39
6. Burks - Excessive Aggressiveness	.01 (NS)	39	-.05 (NS)	39	.19 (NS)	39	.07 (NS)	39

(1) All correlation coefficients not marked are significant at the .01 level.

Table 5, above, suggests that the social skills being measured by this part of the Five P's would seem similar to those measured on the California Preschool Social Competency Scale. Table 5 also suggests that the sub-scales on this part of the Five P's are also similar to the sub-scales of Rapport and Responsibility on the Preschool Attainment Record. Little supportive evidence was found from the Burks Behavior Rating Scale. In general, Table 5 provides strongest evidence for the peer relation sub-scale, next strongest for adult relation and

only modest support for the Emerging Self sub-scale.

Scale 5 - Cognitive Development:

The Cognitive Development Scale consists of two sub-scales. The first sub-scale is Attitudes to Learning which is composed of twelve (12) items. The second sub-scale is labeled Cognitive Skills and contains 46 items. The total scale has fifty-eight (58) items. Scores for each of the sub-scales were calculated by adding points earned on each item, and a total score for Cognitive Development was calculated by summing the scores of both of the sub-scales. These scores were then correlated with the scores on certain other "external" tests to investigate the validity.

The Cognitive Development Scale was validated against both the Leiter (1952) and the McCarthy (1972), which are standard intelligence tests which seem appropriate for use with the children at PSW. In addition, certain sub-tests of the PAR and the Burks were also used. The Leiter and the McCarthy were administrated by the PSW school psychologist, while the PAR and the Burks were scored by the child's teacher. Table 6 presents the evidence relating to the validity coefficients for the Cognitive Development Scale.

Table 6 - Validity Correlation Coefficients for the Cognitive Development Scale Sub-scales on the Five P s (1)

External Scales	Attitude to Learning	N	Cognitive Skills	N	Total Cognitive Development	N
1. Leiter-M.A.	.32 (NS)	27	.49	27	.48	27
2. McCarthy-Imitative Action	.52	29	.59	29	.61	29
3. PAR - Ideation	.45	41	.71	41	.70	41
4. PAR - Rapport	.58	41	.56	41	.51	41
5. Burks-Poor Intellectuality	-.50	39	-.41	39	-.45	39

(1) All correlations were significant at the .01 level except those designated as (NS), not significant.

Table 6, above, suggests that the Cognitive Skills sub-scale has the strongest evidence of validity, yielding significant coefficients of .49 and .59 with established intelligence scales, plus very high coefficients, .71, on the Ideation sub-scale of the PAR. The evidence in support of the sub-scale named Attitude to Learning presents moderately strong coefficients, but it should be understood that the nature of this sub-scale is partly attitudinal, and no extremely high coefficients should be expected with measures of intellectual ability such as the Leiter. It is hoped that further factor

analytic studies will more clearly define this Attitude to Learning sub-scale.

SUMMARY AND CONCLUSIONS

In summary, fourteen (14) sub-scales were validated in this study using a randomized stratified sample of forty-one (41) children. In the case of seven sub-scales it was possible to establish concurrent validity by the application of already established scales, which were applied by personnel other than the teachers in the classroom. These scales, Motor Development, Language Development, and Cognitive Development produced validity coefficients in the .50 to .75 range. Many research and statistics books (Kerlinger, 1974, Guilford, 1965) suggest that validity coefficients can be expected between .00 and .80. The validity coefficients which were produced in this study meet those standards and suggest a satisfactory level of "concurrent" validity.

The two scales, Routines and Self-Help and Social Development were validated against other established teacher rating scales. Here again, the validity coefficients were in the .50 to .70 range, which provides moderately good evidence of the "concurrent" validity of these scales. Obviously, the fact that those scales were validated against other teacher

rating scales weakens the case somewhat. The constructs operating in those scales are classroom skills, and it will be very hard to validate those scales using ratings by people outside of the classroom. Perhaps it is not fair to expect that form of validation with scales of this kind. At the same time, it is hoped that the further factor analyses planned for 1983 will allow us to be more certain about the validity of those two scales: Routines and Self-Help and Social Development. We would hope that some strong "construct" validation evidence will emerge from the factor analyses.

The major focus of these validation studies which are reported in this paper give substantial evidence to the validity of the Five P's as a teacher observational instrument. It is asserted that the Five P's is a valid observational instrument which can and does measure very important aspects of disabled children's classroom behavior. It measures a wide spectrum of children's behavior, as seen by the classroom teacher. The behaviors are not only from a wide spectrum; they approximate in many instances already established constructs in the field of child psychology. Those measurements carry with them the expected degree of validity which the psychological community expects of instrumentation. It is for these reasons that we present this scale to the profession as a valid and useful new tool in the field of preschool exceptional children.

Reference

Arthur, Grace. Arthur Adaptation of the Leiter International Performance Scale. Chicago: Stoelting Company, 1952.

Beery, K. and Buktenica, N. A. Developmental Test of Visual-Motor Integration, Chicago: Follett Publishing Co., 1967.

Burks, Harold F. Burks' Behavior Rating Scales, Los Angeles; Western Psychological Services, 1981.

Bzoch, Kenneth R. and League, R. Bzoch-League Receptive-Expressive Emergent Language Scale, (REEL) for the Measurement of Language Skills in Infancy. Baltimore: University Park Press, 1970.

Carrow, Elizabeth. Test for Auditory Comprehension of Language. Boston: Teaching Resources, 1973

Doll, Edgar A. Preschool Attainment Record. Circle Pines, Minn., American Guidance Service, 1966.

Guilford, J.P. Fundamental Statistics in Psychology and Education. New York: McGraw-Hill, 1965.

Harman, Harry H. Modern Factor Analysis. Chicago: University of Chicago Press, 1976

Kerlinger, Fred N. Foundations of Behavioral Research. New York: Holt, Rinehart and Winston, 1974.

Levine, Samuel, Elzey, F.F. and Lewis, M. California Preschool Social Competency Scale. Palo Alto, California: Consulting Psychologists Press, 1969

McCarthy, D. Manual for the McCarthy Scales of Children's Abilities. New York: The Psychological Corporation, 1972.

Nunnally, Jum. Psychometric Theory. New York: McGraw-Hill, 1978.